Article:

Kendrick, Kobin H. orcid.org/0000-0002-6656-1439 and Torreira, Francisco (2015) The Timing and Construction of Preference : A Quantitative Study. Discourse Processes. pp. 255-289. ISSN 0163-853X

https://doi.org/10.1080/0163853X.2014.955997

The Timing and Construction of Preference: A Quantitative Study

Kobin H. Kendrick

Francisco Torreira

Max Planck Institute for Psycholinguistics

Corresponding author:

Kobin H. Kendrick

Max Planck Institute for Psycholinguistics

Wundtlaan 1

6525 XD Nijmegen

The Netherlands

Phone: +31 24 3521271

Email: kobin.kendrick@mpi.nl

Abstract

Conversation-analytic research has shown that the timing and construction of preferred responding actions (e.g., acceptances) differ from that of dispreferreds responding actions (e.g., rejections). This could serve as a mechanism for early response prediction. We examined 195 preferred and dispreferred responding actions in telephone corpora and found that the timing of the most frequent cases did not differ systematically. However, after approximately 700 ms, the proportion of dispreferreds was greater than that of preferreds, suggesting that a gap of 700 ms, but not shorter, could allow one to predict the responding action. We also examined the timing of dispreferred turn formats (i.e., those with forms of qualification) as an orthogonal dimension to action. Small departures from a normal gap duration (>300 ms) decrease the likelihood of an unqualified acceptance and increase the likelihood that a response, be it acceptance or rejection, will have a dispreferred turn format.

The Timing and Construction of Preference: A Quantitative Study

In conversation each turn at talk presents a potential next speaker with a set of contingencies that must be managed within a very short period of time. A next speaker must, at a minimum, recognize the meaning and action of the turn and prepare a relevant next turn. But, as Levinson (2013) points out, the gaps between turns, averaging between 200 to 300 ms, do not provide sufficient time to plan even a simple utterance, which psycholinguistic research has shown requires at least 600 ms (Indefrey & Levelt, 2004). In order to produce a relevant next turn within such a short period of time, a next speaker must therefore anticipate or project at least part of the current turn (Sacks, Schegloff, & Jefferson, 1974). In light of such constraints, the beginning of a turn at talk – that is, its first verbal components – becomes a crucial resource for projection, enabling a next speaker to plan his or her turn in advance (Schegloff, 1987; Levinson, 2013). Yet even before such initial components, the timing of a turn's onset may also facilitate early projection. Conversation analysts have suggested that the time it takes for a recipient to respond to a question can foretell the valence of the answer (see, e.g., Clayman, 2002). This would, in principle, enable a next speaker to begin to plan a next turn at the earliest possible moment, even before the current turn has officially begun.

Here we report on a series of quantitative conversation-analytic investigations into a mechanism that could enable early projection of a turn on the basis of timing alone. This mechanism, first described qualitatively in the conversation-analytic literature under the rubric of preference organization (e.g., Schegloff, 1988), depends on a systematic distinction between alternative response types, referred to as preferred and dispreferred responses (e.g., accepting or rejecting a request, offer, or invitation). Research has shown that preferred responses come quickly and take simple forms, whereas dispreferred responses occur after significant delay and exhibit more complex constructions, often with prefaces, qualifications, or accounts (see Background). Thus the timing of a response alone may signal whether it will be preferred or dispreferred, one solution to the psycholinguistic puzzle of turn-taking.

The weight of these implications places a heavy burden of proof on generalizations about the timing of preferred and dispreferred responses. The conversation-analytic literature offers a wealth of qualitative evidence, in the form of single cases and small collections of cases, that supports these generalizations (see, e.g., Atkinson & Drew, 1979; Pomerantz, 1984; Heritage 1984), and results from one large quantitative study (Stivers et al., 2009) have also been cited as support (Pomerantz & Heritage, 2013). But the claim that the timing of preferred and dispreferred responses differ systematically, such that timing alone could serve as a reliable signal of the responding action, has not been verified. The aim of the current study is thus twofold: first, to reproduce and specify the classic finding that dispreferred responses tend to be delayed and, second, to evaluate the claim that the timing of a response could enable one to predict its status as preferred or dispreferred.

Background

The basic insight of research on *preference* in conversation analysis (CA) is that the practices speakers use in interaction exhibit systematic asymmetries that serve to maximize opportunities for affiliative actions and minimize opportunities for disaffilative ones (Heritage, 1984). The concept of preference thus does not refer to the psychological states of the speakers, but rather to socially normative principles that speakers observably orient to in interaction (Pomerantz & Heritage, 2013). Research on preference has examined the practices that speakers use to respond to a variety of different actions, including compliments (Pomerantz, 1978), self-praise (Speer, 2012), and complaints (Dersley & Wootton, 2000), as well as those that speakers use to initiate sequences of action (Robinson & Bolden, 2010). In this section, we review the literature on preference that concerns (i) the timing and construction of preferred and dispreferred responding actions, (ii) a distinction between the action a response performs and the form or format that a response takes, and (iii) the claim that delay before a response may signal that a dispreferred action is likely.

A basic observation in the literature on preference is that responding actions that align with (accept, agree with, grant, etc.) an initiating action tend to take a different form than those that fail to align (reject, disagree with, deny, etc.). This observation has frequently been illustrated through a comparison of two paradigm cases drawn from the same telephone call, presented here in Extracts (1) and (2) (Atkinson & Drew, 1979, p. 58; Levinson, 1983, p. 333-4; Heritage, 1984, p. 265-6; Clayman, 2002, p. 233; Raymond, 2003, p. 943).

```
(1) SBL 1-10
01 Ros:     Why don’t you come'n see me so:me[ti:mes.
02 Bea:                                      [I would li:ke to:.
(2) SBL 1-10
01 Ros:     And uh the: if you'd care tuh come ovuh, en visit u
02          little while this morning I'll give you cup of coffee.
03 Bea:     Uhh-huh hh Well that’s awfully sweet of you I don't
04          think I can make it this morning, hheeuhh uh:m (0.3)
05          .tch I’m running an a:d in the paper 'nd an:d uh hh I
06          I haveta stay near the pho::ne,
```

In both cases, Rose extends an invitation to Bea, a first pair part (FPP) of an adjacency pair sequence that makes acceptance or rejection conditionally relevant as a second pair part (SPP). That is, the relevance of these responding actions arises only upon the recognizable production of the invitation (Schegloff & Sacks, 1973; Schegloff, 2007). In the first case, Bea accepts the invitation, whereas in the second she rejects it. In response to an invitation, rejection is a *dispreferred action* insofar as it fails to align with the action of the first pair part, does not support the accomplishment of the activity, and threatens social solidarity

(Heritage, 1984; Schegloff, 2007). Acceptance, which takes the opposite stance towards the invitation, is a *preferred action*.

A comparison of the construction of the two actions shows that the rejection employs a number of practices that the acceptance does not: (i) prefatory particles and in-breaths; (ii) qualification and mitigation; (iii) appreciation of the invitation; (iv) placement of the rejection component in a non-contiguous position with the invitation; and (v) an account that explains the speaker's inability to accept the invitation (see Atkinson & Drew, 1979, p. 58; Levinson, 1983, p. 334-5; Heritage, 1984, p. 265-6). As Heritage (1984) argues, the different formats that Bea employs here to deliver her acceptance and rejection are "characteristic of the general ways in which acceptances and rejections of invitations are accomplished" (p. 266). A responding action that employs turn-constructional practices such as these has a *dispreferred turn format*, whereas a responding action that does not has a *preferred turn format*. One outcome of such practices is that they delay the onset of the base turn-constructional unit (TCU), that is, the word, phrase, or clause that a speaker uses to build his or her turn at talk (Sacks et al., 1974).

A clear distinction between a responding action and the turn format a speaker employs to deliver this action is necessary because, as Schegloff (1988) observes, dispreferred actions can be produced with preferred turn formats and preferred actions can be produced with dispreferred turn formats (cf. Lerner, 1996, p. 305; Heritage, 1984, p. 267-268). The disassociation of actions from the practices that speakers use to construct them implies a two-dimensional possibility space, with action and turn format as orthogonal dimensions (see also Lee, 2013; Lindström & Sorjonen, 2013).

A further generalization that is often made in the literature concerns timing: preferred actions tend to occur relatively early (e.g., in overlap, as in Extract 1) whereas dispreferred actions tend to occur relatively late (see, e.g., Atkinson & Drew, 1979, p. 58; Levinson, 1983,

p. 334). In addition to initial components that can delay the production of an explicit rejection component, silence prior to the initiation of a responding turn can also serve to delay the delivery of a dispreferred action. Schegloff (2007) illustrates this phenomenon with the following cases, in Extracts (3) and (4) below.

(3) NB IV:10 (Schegloff, 2007, p. 68)[1]

01 Lot: ↑Don’t you want me to come dow:n an’ get you
02 tomorrow an’ take you down to the beauty parlor?
03 -> (0.3)
04 Emm: What fo:r I ↑just did my hair it looks like proh-
05 a profess↓ional.

(4) Erhardt 1 (Schegloff, 2007, p. 68)

01 Kar: °Gee I feel like a real nerd°<You can all come up here,
02 -> (0.3)
03 Vic: Nah, that’s alright we’ll stay down here,

In each case, a gap of approximately 300 ms occurs prior to the beginning of a dispreferred action. For Schegloff (2007), these cases are instances of a general tendency in the timing of preferred and dispreferred actions: “The transition space between the first pair part turn and a dispreferred second pair part turn is commonly overlong” (p. 67). This parallels the observation by Heritage (1984) that “acceptances to offers, invitations, etc. commonly occur ‘early’, i.e., immediately on completion or in slight overlap with their first pair parts” whereas “rejecting responses very often occur ‘late’” (p. 273). In a recent review, Pomerantz and Heritage (2013) state this even more plainly: “rejections are performed with delays before turn initiation” (p. 210). In general, then, the literature describes a difference in the timing of preferred and dispreferred actions as a one of relative frequency. In the words of

[1] In the Background and Data and Methods sections, all transcripts preserve the timings in the original sources, although other aspects of the transcripts may have been altered to improve readability.

Heritage (1984), these differences are "strongly recurrent and patterned design features of these two classes of actions" (p. 274).

A quantitative study on the timing of responses to polar questions by Stivers et al. (2009) lends support to this generalization. In a sample of 219 polar questions, Stivers et al. observe that affirmations (e.g., "yeah") tend to occur faster than disaffirmation (e.g., "no"), with mean gap durations of approximately 75 ms and 400 ms, respectively. Although Stivers et al. do not address the matter of preference directly, the results demonstrate that the timing of response alternatives can indeed differ, in line with previous findings. However, because the study reports only the mean gap durations and does not show the full distribution of the data, one cannot determine whether the timing of the two alternatives differs systematically (i.e., whether the most frequent disaffirmations occur later than the most frequent affirmations) or whether the difference in means is driven by a minority of cases. The systematicity of the timing of response alternatives is crucial to the question of whether timing alone could serve as a mechanism for early projection. We return to this issue in the discussion of the current study.

Since the initial discovery of the phenomenon (Sacks, 1973/1987), the difference in the timing of preferred and dispreferred responding actions has been understood not merely as an empirical fact, but also as a meaningful signal that a second pair part is likely to be a dispreferred alternative. The sequence in Extract (5) below has been presented to illustrate this point (see Sacks, 1973/1987, p. 64; Pomerantz, 1984, p. 76-7; Schegloff, 2007, p. 71; Pomerantz & Heritage, 2013, p. 214).

```
(5) JS:II:48 (Pomerantz 1984:77)
01 A:     D'they have a good cook there?
02        (1.7)
03 A:     Nothing special?
04 B:     No. Every- everybody takes their turns.
```

The design of the initial question (line 1) prefers a response that affirms the proposition. After a gap of 1700 ms, however, the speaker reformulates the question, offering a candidate answer (line 3), in a form that reverses the preference of the initial version, allowing the recipient to produce "no" as a preferred action. The speaker apparently understands the 1700 ms of silence as indicative of a dispreferred action and uses this information to revise the initial question. On the basis of cases such as this, Clayman (2002) concludes that "interactants can anticipate the type of response that is forthcoming purely on the basis of its initial form. . . . any delay in responding – even mere silence – may be interpreted as the first move toward some form of disagreement/rejection" (p. 235).

The interpretation of silence in conversation has also been the focus of experimental research. Roberts, Francis, and Morgan (2006) asked participants to listen to simulated telephone calls in which speakers accepted requests (e.g., A: "Can you give me a ride over there?" B: "Sure!"), with the gaps before the acceptances manipulated to be 0, 600, or 1200 ms long. Participants then rated the 'willingness' of the speakers of the acceptances in each condition on a six-point scale, answering questions like "How willing is Rachel to give her friend a ride?". The results showed that ratings decreased as gap durations increased, indicating that subjects monitored the gaps and used this information to make attributions of willingness (see Roberts, Margutti, & Takano, 2011 for a cross-linguistic replication). Drawing on observations from CA, the authors argued that delays before acceptances signal "interactional trouble" and that this, in turn, influenced participants' ratings of willingness. While the results do not directly bear on the question of whether recipients can use timing alone to predict the incipient action, in that the judgments were made after participants had heard the acceptances, they do demonstrate that listeners are sensitive to gap durations and can use this information to make social attributions.

Current Study

The first goal of the current study is to verify and further specify observations in the CA literature on the timing and construction of preferred and dispreferred responding actions through a quantitative analysis of a large collection of cases. With respect to timing, one clear prediction that we can derive from the literature is that the duration of the gap between the end of a first pair part and the beginning of a second pair part should be systematically greater for dispreferred actions than for preferred actions. If the difference is indeed strongly recurrent, one expects to observe the difference not only in a small number of extreme cases but also in the most frequent cases as well. With respect to turn construction, a prediction based on observations in the literature is that, in comparison to preferred actions, dispreferred actions should include more turn-initial practices that delay the production of the base turn-constructional unit.

In addition to predictions such as these, the literature also points to an open question. The analysis of preference as both an action-based and a format-based phenomenon raises the question of whether the claims in the literature hold equally under both analyses. If indeed a dispreferred action can be done in a preferred turn format, or vice versa, then the question becomes whether the generalizations about timing pertain to dispreferred actions per se, to dispreferred formats, or to both.

The second goal of the current study is to evaluate the claim that the timing of a responding action can serve as an early signal of preference and can therefore enable a recipient to project whether an incipient response is likely to be acceptance or rejection. This claim also raises a number of questions. If even mere silence can signal that rejection is imminent, one may ask *how much* silence would a recipient need to hear in order to make this prediction? Extracts (3) and (4) suggest 300 ms may be sufficient, whereas Extract 5 points to larger values. The evaluation of this claim depends on how one measures the gap between first and second pair parts. An examination of the sequence in Extract (2) reveals that an estimate of the timing of the response depends greatly on the method of measurement. If one measures from the end of the first turn to the first audible component of the response, then the response comes quickly (after only 151 ms). However, if one excludes breathing and prefatory particles and measures to the first word of the first base turn-constructional unit (i.e., "that's"), then one would conclude that the response comes after a significant delay (767 ms).

To answer these questions, the current study systematically analyzes a large sample of responses to invitations, offers, requests, suggestions, and proposals, exhaustively drawn from corpora of telephone calls that have been widely used in CA research. After we describe the data and methods and present the frequencies of preferred and dispreferred actions and turn-initial practices, we present the results of two main investigations. First, we report on an analysis of the timing of preferred and dispreferred responding actions, based on three different measures of the temporal offset between the first and second pair parts. Second, because these results do not straightforwardly conform to the predictions outlined above, we report on a subsequent analysis of the timing of preferred and dispreferred turn formats, in which we distinguish acceptances that qualify a speaker's commitment to the course of action and explicit rejections without qualification or mitigation.

The results of these investigations depend crucially on the methods we employ to measure the timing of turn-taking. To test whether our results could be an artifact of our objective method of timing, we also report on the relationship between rate of speech and timing and compare our measurements to those of Gail Jefferson in her transcriptions of a subset of the recordings. Finally, in light of the results of the study, we reassess the generalization in the CA literature that dispreferred actions tend to be delayed and consider the reliability of timing as an early signal of preference.

Data and Methods

Data

The data for this study come from corpora of telephone calls: SBL, NB, Holt, XTR/YYZ, Kamunsky, HGII, SF, and TG. A total of 185 calls were examined, within which 238 sequences were systematically identified. For a sequence to be included, the FPP had to be recognizable as a request, offer, invitation, proposal, or suggestion. These actions form a natural class in that (i) each makes relevant an SPP in which a speaker either commits or fails to commit to a future course of action and (ii) each prefers acceptance over rejection as responding actions.[2] To identify these actions, we drew on the relevant CA literature on requests (Curl & Drew, 2008), offers (Curl, 2006), invitations (Drew, 1984), and proposals (Maynard, 1984; Houtkoop-Steenstra, 1987, 1990). Couper-Kuhlen's (in press) study of the grammatical formats of requests, offers, suggestions, and proposals was especially helpful. Eleven sequences were excluded from the analysis because the beginning of the SPP could not be precisely identified due to overlap (see below), and 33 sequences were excluded because they included increments or insert sequences (see below). The final collection included 195 sequences.

Increments and Insert Sequences

Because a central concern of the study is the timing of gaps measured in milliseconds, two types of cases were systematically excluded: (i) sequences in which the FPP included an increment after a gap or pause and (ii) those in which the SPP occurred after an insert

---

[2] A possible exception to this generalization is offers. Schegloff (2007) observes that "although generally it appears that accepting is the preferred response to offers. . .this may be contingent on the item being offered and the context." Because this has not yet been subject to systematic investigation, we operate under the assumption that offers prefer acceptance over rejection (see Heritage 1984, p. 269).

sequence. Extracts (6) and (7) illustrate the two types, respectively.

```
(6) Holt U88:1:8
01 Gor:     .tch Are you gonna drive in. Cz I n- I know there
02          wz some rumor about it,
03          (0.5)
04 Gor:     .hhhh Or not.
05          (0.5)
06 Dan:     No but I'll be downtown (0.2) at nine forty five.
(7) Holt SO88:1:5
01 Gor:     Well uh- (0.2) How 'bout t'morrow night.
02          (0.6)
03 Gor:     .ts.kh
04          (0.5)
05 Sus:     T'morrow ni:gh[t.
06 Gor:                   [Sat'dee. .t.k
07 Sus:     .h I don'know about that, .h
```

Although increments and insert sequences are relevant to the study of preference (Pomerantz, 1984; Davidson, 1984), they also obscure the measurement of the durations between FPPs and SPPs. All such cases were excluded from the study (*n*=33).

Analysis of Preference

In an effort to maximize the replicability of our results, we employed a straightforward and objective method for the analysis of preferred and dispreferred actions: the next turn after the FPP was analyzed as the SPP. If the SPP included a turn component that accepted (agreed with, confirmed, granted, etc.) the FPP, regardless of its position in the turn and regardless of other components in the turn, we analyzed it as a preferred action. If the SPP did not have such a component, we analyzed it as a dispreferred action. This method has the advantage that it does not rely on the presence or absence of turn-initial particles, in-breaths, or the duration of the transition space, the features whose distribution we wish to test.

A variety of turn formats were thus analyzed as preferred actions. These included acceptances done with single lexical items (e.g., "Mkay."), those done with multiple lexical items (e.g., ""↑YES. ↓Su:re."), positive assessments (e.g., "Brilliant."), repetitional confirmations (e.g., ".pt u-Will do:."), acceptance components in non-initial position (e.g., ".hhhhhh >Sure.<"), and acceptance components with qualification or mitigation (e.g., "Ye:s I think so:,"). In a minority of cases, an acceptance component occurred together with one or more other components that worked to mitigate, qualify, or otherwise undermine the acceptance. That is, the turn employed a dispreferred turn format (see Background), as in the following cases.

```
(8) Holt O88:1:9
01 Ed:      .kh What about Thursday ev[eni n g.]
02 Les:                              [ihYe- e-]↑Yes: e-u-
03          ↑Thursday: .hhh (0.2) Oh:.↑Thursdee I'm f:::-
04          I'm going ou:t about half pas' sev'n. But before
05          that I'm in.
(9) NB IV:10
01 Emm:     ↑If you wanna dri:ve down'n see me ah'd love duh
02          see yuh,
03 Lot:     .hhh O:kay well I gottle a few thing
```

In each case, the second pair part includes an acceptance component, which is either followed by a specific qualification (Extract (8)) or an inability account (Extract (9)). Such cases were analyzed as preferred actions due to the presence of acceptance components. We consider dispreferred turn formats in detail in a later section (see Qualified Acceptances).

All SPPs that did not include an acceptance component were analyzed as dispreferred actions. In contrast to acceptances, which frequently include explicit acceptance components, rejections normally do not include explicit rejection components such as "no" (Kitzinger & Firth, 1999). Furthermore, the recognizable withholding of acceptance can be understood by participants as tantamount to rejection. The following cases were therefore analyzed as dispreferred actions.

```
(10) NB II.2
01 Emm:      W'l  GIVE ME A BU:ZZ if you u (0.2) uh c'm o:n
02           down if you eh ah'd li:ke tih have yuh come do:n
03           fe[r a]
04 Nan: ->     [Yer] a r'l ↑sweetheart.
05           (0.7)
06 Nan:      Wish you hadda car so you c'd c'm over he:re,
(11) SF 1
01 Mar:      Why: don:'t if I don't see you en I probably will
02           b'fore Saturday¿ (0.8) .t .khhhhh uh:m why don'tchu
03           call me Saturday morning.
04 Joa: ->   u-W'l ah'll probly call you Thursday.
```

In response to Emma's invitation in Extract (10), Nancy produces an appreciation that occupies the position in which an acceptance could have been done. In Extract (11), rather than accept Mark's proposal, Joanne counters with an alternative proposal, which in this context amounts to a rejection. We examine dispreferred actions at a finer level of granularity in a later section (see Flat Rejections). To assess the reliability of this analysis, a research assistant with no prior experience in CA coded 20 percent of our sequences using the criteria above. The vast majority of cases (87.2%) received the same analysis as our own, yielding a substantial level of agreement in a Cohen's kappa coefficient ($\kappa = 0.69$).

Measurement of Timing

The beginning of a turn at talk is a complex phenomenon (see Schegloff 1996). In order to account for the complexity of turn beginnings, a series of temporal offsets were calculated on the basis of four recognizable points in the sequence. These points are illustrated in Figure 1 for the sequence in Extract (12) below.

Point 1 The last acoustic signal attributable to the articulation of the FPP. Vocal noises (e.g., out-breaths, clicks) at the end of the turn were ignored.

Point 2 The first acoustic signal attributable to the SPP, whether a click, in-breath, prefatory particle, or the first word of a base TCU.

Point 3 The beginning of the first particle or word of the SPP, after turn-initial clicks and in-breaths.

Point 4 The beginning of the first word of the base TCU, after turn-initial in-breaths, clicks, the prefatory particles “well” and “u(h)m”.

```
(12) Holt O88:1:9
01 Les:     Couldn't you adverti:se amongst teachers a bit,
02 Ed:      .hhh °Ahh::m::° (0.6) Yes I spoze I cou:ld.
```

These points were located manually through the inspection of waveforms and narrow-band spectrograms in Praat. The question of whether this method of timing may be problematic is taken up in a later section (see Are Objective Measurements of Timing a Problem?). They were initially located by the first author and then checked by the second author. Cases of disagreement were examined by both authors together and discussed until the disagreement was resolved. Figure 1 shows the four temporal points in the sequence in Extract (12).

Once the temporal points were identified, three temporal offsets were calculated for each sequence. Cases of overlap between the end of the FPP and the beginning of the SPP were given negative values.

Offset 1 The duration between Point 1 and Point 2. That is, the duration between the end of the FPP and the first audible component of the SPP. This offset is a measure of the gap between the FPP and SPP.

Offset 2 The duration between Point 1 and Point 3. That is, the duration between the end of the FPP and the first word or particle of the SPP. This offset is a measure of the gap together with the pre-beginning phase of the SPP, which includes clicks and in-breaths within the duration of the offset but not prefatory particles such as “u(h)m” or “well”.

Offset 3 The duration between Point 1 and Point 4. That is, the duration between the end of the FPP and the first word of the base TCU. This offset is a measure of the gap together with the pre-beginning and beginning phases of the SPP, which include clicks, in-breaths, and the prefatory particles “u(h)m” and “well”.

Note that the offsets overlap, in that Offset 2 encompasses Offset 1, and Offset 3 encompasses Offsets 1 and 2. To assess the reliability of our timing measures, a research assistant coded 20 percent of our sequences using the criteria above. Very high agreement was achieved for the presence of in-breaths and clicks (κ = 0.87) and prefatory particles (κ = 0.92). For the continuous offset measures, a high correlation between the second annotator’s measures and our own was observed (Offset 1: r = .97; Offset 2: r = .83; Offset 3: r = .99). In 87.2% of the cases, and for all the three offsets, the second annotator’s measures and our own did not differ by more than 100 ms.

Results

The subsequent sections report the results of our investigations into the timing and construction of preference. After we describe the frequency of preferred and dispreferred actions and the occurrence of turn-initial practices in the collection, we examine the timing of preferred and dispreferred actions for three temporal offsets. In light of the unexpected results of this analysis, we next examine the timing of preferred and dispreferred turn formats, in addition to actions. Finally, we consider whether our results could be an artifact of our method of timing.

The Frequency of Preferred and Dispreferred Actions

Within the collection of 195 responses, 63 percent (n=123) were analyzed as preferred actions, with the remaining 37 percent (n=72) analyzed as dispreferred actions. The greater proportion of preferred actions is consistent with previous research on preference. Raymond (2000, p. 106) observes that 75 percent of responses to polar questions (n=243) align with a preference for type-conformity, containing some form of “yes” or “no” (see also Raymond, 2003). Stivers (2010, p. 2778) finds that the vast majority of responses to polar questions (80%; n=183) conform to a preference for answers over non-answers (see Stivers & Robinson, 2006). In our study, however, the proportion of preferred actions is statistically lower than those observed both by Robison ($\chi^2(1) = 7.76$, $p < .001$) and Stivers ($\chi^2(1) = 18.16$, $p < .0001$). One possible explanation for this is that the studies by Raymond and Stivers primarily concern polar questions that request information. Stivers (2010, p. 2776) reports that less than 3 percent of polar questions in her study were suggestions, offers, or requests. Actions that request personal commitments to future courses of action may be inherently more vulnerable to rejection than requests for information.

The Frequency of Turn-Initial Practices

The use of turn-initial practices, such as turn-initial breaths (TIBs) and turn-initial particles (TIPs), has been associated with the construction of dispreferred actions (see Background). As these practices successively delay subsequent phases of a turn, an analysis of their occurrence was a necessary prerequisite to the analysis of the timing of preferred and dispreferred actions reported below. The frequency of TIBs and TIPs differs systematically between preferred and dispreferred actions, in that both occur more frequently in dispreferreds. Nearly half of all dispreferred actions include a TIB (42.3%; n=32), whereas only 15.8 percent (*n*=21) of preferred actions do. As for TIPs, the frequency of non-lexical hesitations particles "u(h)m" is greater in dispreferred actions than in preferred actions (19.4% and 5.7% respectively), and prefatory particle "well" occurs in over one quarter of all dispreferred actions (29.2%) and only 4.9 percent of preferreds. Considering TIBs and TIPs together, we found that the majority of dispreferreds (65.2%) include at least one of these features, whereas only 23.6 percent of preferreds do. The probability that a response will be a dispreferred action given the presence of these features is also worth noting. When a response includes a turn-initial "well", the probability that it will be a dispreferred action is 0.78; when the response includes turn-initial "u(h)m" the probability is 0.67; and when it includes turn-initial in-breath the probability is 0.6. In contrast, when a response lacks any of these practices, the probably that it will be a dispreferred action drops to 0.21. A series of mixed-effects logistic regression models with preference as dependent variable, each of the TIPs and TIB as a fixed predictor, and speaker as random factor, yielded statistically significant effects for each of the fixed predictors (in-breath: $\beta = -1.55$, $z = -3.88$, $p < .0005$; "u(h)m": $\beta = -1.7$, $z = -3.12$, $p < .005$; "well": $\beta = -2.14$, $z = -3.9$, $p < .0001$).

The Timing of Preferred and Dispreferred Actions

This section presents quantitative results for three temporal offsets between first pair parts and second pair parts. The goal of this analysis is, firstly, to verify the generalizations in the literature that dispreferred actions tend to be delayed and, secondly, to evaluate the claim that the timing of a response alone may be a reliable signal of the responding action.

The Timing of the First Audible Component. We first examine the duration of turn-initial silence (Offset 1) for preferred and dispreferred actions. Figure 2 shows a density plot of the distribution of turn-initial silence for preferred and dispreferred actions. An initial observation is that the two distributions have different shapes but overlap substantially within a temporal window of approximately -100 to 500 ms. This suggests that preferred and dispreferred actions do not form two clearly distinct groups. Outside this temporal window, the distributions of preferred and dispreferred actions overlap far less, due to two distinct groups of early and late dispreferred actions. After 750 ms, the proportion of dispreferred actions increases considerably (15.3% of dispreferreds vs. 4.1% of preferreds; $\chi^2(1) = 6.17$, $p < .05$). A greater proportion of responses that begin prior to -100 ms are also dispreferred actions, which suggests that responses that begin very early may be more likely to be dispreferred actions, although, contrary to previous case, this trend fails to reach statistical significance, perhaps due to the small number of observations in this range ($\chi^2(1) = 2.5$, $p = .11$). Within the temporal window of -100 ms to 700 ms, which includes the bulk of the data for both preference groups, dispreferred actions tend to be slightly earlier than preferreds. This difference was statistically significant in a mixed-effects regression model with Offset 1 as the dependent variable, preference as a fixed factor, and speaker as a random factor ($\beta = 63.1$, $t = 2.01$, $p < .05$). This observation runs counter to the claim by Schegloff (2007, p. 67) that the transition space before dispreferred actions is “commonly overlong”.

Qualitative analysis of individual cases points to the relatively high frequency of turn-initial in-breaths, which occur in close to half of all dispreferred actions, as one factor that influences the duration of the transition space. This effectively pulls the distribution of dispreferred actions forward in time, closer to the end of the FPP, since preparatory breathing often occurs in overlap with the end of the prior turn.[3]

---

[3] Unlike turns at talk, which conform to a ‘one speaker at a time’ constraint (Sacks et al., 1974), some forms of breathing freely co-occur in overlap with another speaker’s talk. Insofar as current speakers do not employ resources for the management of overlapping talk (Schegloff, 2000), they do not orient to recipients’ breathing as competitive with a current turn. This suggests that breathing is not ‘turn-organized’ and is therefore not subject to the same constraints on timing as turns. That said, as Schegloff (1996) observes, a turn-initial in-breath is a preparatory action that can signal an intention to speak.

The Timing of the First Word or Particle. Figure 3 shows a density plot of the distribution of Offset 2 for preferred and dispreferred actions. A comparison with the results for Offset 1 yields three important observations. Firstly, since preferred actions rarely start with turn-initial in-breaths or clicks (only 17.1% do), their distribution is very similar to that shown for Offset 1. In contrast, the distribution of dispreferred actions, which often start with these pre-beginning components (44.5%), clearly differs between Offsets 1 and 2. Secondly, the mode for dispreferreds for Offset 2 occurs later than for Offset 1, at around 250 ms, and is roughly equal to the mode for preferreds. The fact that the mode of dispreferreds occurs earlier than that of preferreds for Offset 1 can be therefore attributed to the high frequency of pre-beginning components such as in-breaths and clicks. Thirdly, the distribution of Offset 2 for dispreferred actions shows significantly more variability than the distribution of Offset 1 for this same group of responses. While Offset 1 values were concentrated within a window of -100 to 500 ms, Offset 2 values tend to extend beyond 500 ms more often. This greater variability can again be partly attributed to the occurrence of pre-beginning components in dispreferred actions which displace the beginning of the turn to a variable extent, depending on the duration of the in-breaths, clicks, and post-pre-beginning silences.

The results for Offset 2 also reveal that the proportion of dispreferred actions is greater after approximately 700 ms (25% of dispreferreds vs. 8.9% of preferreds; $\chi^2(1) = 8.02$, $p < .005$), but before 0 ms the proportion of preferred actions is only slightly greater (9.7% of dispreferreds vs. 12.2% of preferreds; $\chi^2(1) = 0.08$, $p = .77$). The majority of all responses, preferred and dispreferred, fall within the same temporal window between roughly 0 and 700 ms.

The analysis of Offset 2 indicates that the unexpected observation that dispreferred actions occur earlier than preferreds, in terms of the modes of the distributions, is the result of the high frequency and early onset of the pre-beginning phase of dispreferred actions. Yet

Offset 2 also yields its own unexpected observation: the timing of the most frequent preferred and dispreferred actions is virtually the same.

The Timing of the Base Turn-Constructional Unit. Figure 4 shows a density plot for preferred and dispreferred actions for Offset 3. Although the distribution of preferred actions for Offset 3 is very similar to the previous measures, the distribution of dispreferred actions clearly differs, showing significantly more variability than for Offset 1 and 2. After 600 ms, the proportion of dispreferreds is much greater than that of preferreds (47.2% of dispreferreds vs. 19.5% of preferreds; $\chi^2(1) = 16.56$, $p < .0001$), which suggests that after 600 ms responses are more likely to be dispreferred actions. This is because turn-initial practices that delay the onset of the base TCU are much more frequent in dispreferreds (see The Frequency of Turn-Initial Practices). Before -100 ms, the proportion of preferreds and dispreferreds is similar (6.9% of dispreferreds vs. 4.9% of preferreds; $\chi^2(1) = 0.11$, $p = .74$). Like for the previous measures, the bulk of all responses, both preferred and dispreferred actions, fall within the same temporal window, which here spans from approximately 0 to 800 ms. In comparison to Offsets 1 and 2, however, the mode for dispreferreds for Offset 3 occurs much later, at roughly 600 ms, and is significantly later than the mode for preferreds, which occurs close to 300 ms. Along the same lines, the difference of means between the two groups is statistically significant in a mixed-effects model with Offset 3 as dependent variable, preference as a fixed factor and speaker as a random factor ($\beta = -265.3$, $t = -3.23$, $p < .005$).

Summary. The results of the analysis of the timing of preferred and dispreferred actions did not provide clear evidence that speakers systematically delay the onset of dispreferreds responding actions. Across all three measures of timing, we observed that a substantial proportion of dispreferred actions occurred relatively early, between 0 and 400 ms, and thus did not differ systematically from preferreds. We also observed, however, that dispreferred actions were far more likely to include turn-initial practices, such as in-breaths and particles, that successively delay the onset of the subsequent phases of the response. But these turn-initial practices do not expand the duration of the transition space per se (i.e., Offset 1) and in the case of turn-initial breath can in fact contract it. The results also clearly demonstrate that, regardless of the measure of timing, responding actions that occur very late, after approximately 700 to 800 ms, are with very few exceptions dispreferreds. Thus while we do not find that long delays are characteristic of dispreferred actions, since short gaps are the most frequent before dispreferreds, we do find that dispreferred actions are characteristic of long delays.

The Timing of Preferred and Dispreferred Formats

The results of the analysis in the previous section show that it is unexpectedly common for preferred actions to occur after relatively long delays and for dispreferred actions to occur after little or no delay. This comes as a surprise in light of the generalizations in the literature and warrants explanation. One possible explanation, which we pursue in this section, is that the timing of a response is not a feature of the responding action per se (i.e., not a feature of doing acceptance or rejection), but rather a feature of the turn's construction, one which can occur with acceptances and rejections alike. That is, a delay before the initiation of a responding action may be a feature of a dispreferred turn format, not a dispreferred action (see Background).

To evaluate this possibility, in this section we add an analysis of turn format to our previous analysis of preferred and dispreferred actions. We show that preferred actions can employ dispreferred turn formats, which we refer to as *qualified acceptances*, and equally that dispreferred actions can employ preferred turn formats, which we call *flat rejections*. The analysis of turn format together with the previous analysis of preferred and dispreferred actions results in four response types, given in Table 1.

In the remainder of this section, we first illustrate and formally define qualified acceptances and flat rejections and then examine the timing of all four response types individually. Finally we consider the timing of preferred and dispreferred formats and compare this to the timing of preferred and dispreferred actions.

Qualified Acceptances. Although preferred actions typically employ relatively simple turn formats (e.g., a straightforward and unqualified "yeah"), they can also have more complex designs that qualify or mitigate the speaker's commitment to the action and thereby formally resemble dispreferreds. Here we present and analyze a number of such cases in which preferred actions are delivered through dispreferred turn formats and then introduce a set of formal criteria that we use to identify these *qualified acceptances* within the data.

An initial observation is that some preferred and dispreferred actions employ a turn format that, in one form or another, qualifies the speaker's commitment to the action. Consider the two responses below, an acceptance and a rejection, respectively.

```
(13) Holt 10:88:1:11[4]
01 Ski:     Uh:m (.) would Sundee be alri:ght.h.h
02          (563 ms)
03 Joy:     ehYe:s as far as I: kno:w?
(14) 3a A&B1
01 Bel:     Do you need any last minute things?
02          (878 ms)
03 Anne:    Uhm:: (0.9) .mhh (0.2) no I don't think so.hh
```

In the first case, Skip has called Joyce to confirm the date of an invitation and at line 1 proposes Sunday. The format of Joyce's response includes a brief non-lexical hesitation "eh", an acceptance token "Ye:s", and the qualification "as far as I: kno:w?" with which she treats her acceptance as contingent upon that of a third party. (Indeed, in her next turn, not shown,

[4] The transcripts in this section present our own timing measurements, in milliseconds, for all relevant gaps.

she confirms the date with her husband.) In the second case, after Bel has called Anne to offer to help out for a dinner at Anne's house later that day, Anne rejects the offer with the turn format that includes a prefatory particle "Uhm::", a rejection token "no", and a component that qualifies or mitigates this rejection "I don't think so". Thus while the responding actions differ across these two cases, the turns that the speakers use to deliver them have similar formats, in that both include turn components that qualify the speaker's commitment.

The precise phonetic realization of an explicit acceptance component and the informing of a possible conflict can also serve to qualify the speaker's commitment to an acceptance. In the next case, Nina has called Anne to ask her for help preparing for a job interview. After Anne agrees, she proposes that they meet on Monday.

(15) 2b A&N 2

01 Anne: Let's meet on Monday.hh

02 **(500 ms)**

03 Nina: °Okay°.

04 Anne: Come over here in the morning,

05 **(495 ms)**

06 Nina: Okay.=[I have a lunch on Monday.

07 Anne: [We'll (have)

08 Anne: Oh.=Okay.

09 Nina: I have to go to. So maybe after tha:t?

10 Anne: Okay.

11 Nina: At like one [is that okay?

12 Anne: [.hhhhhh >Sure.<

Nina responds with a *sotto voce* "okay" (line 3), which Anne treats as an adequate acceptance, proceeding to the next step in the arrangements (see Houtkoop-Steenstra, 1987). Nina again responds with "okay", which Anne again appears to treat as adequate (line 7), but Nina now qualifies her acceptance through an informing of a possible conflict ("I have a lunch on Monday"; line 6). While the responses that Nina provides accept Anne's proposals and can therefore be analyzed as preferred actions, the acceptance at line 6 can also be understood as less than full, insofar as the speaker qualifies it through an additional turn-unit and the *sotto voce* acceptance at line 3 arguably foreshadows the minor complication that subsequently emerges in the arrangements. Thus the precise construction of the speaker's acceptances in this case signals dispreference, despite the explicit acceptance component (see also Roberts et al., 2006 on the prosody of acceptances).

The response in the next example also accepts the first speaker's proposal but similarly includes turn components that qualify the speaker's commitment to that action. Margy has called Emma to ask her to come help out with some bookkeeping. After Emma tells Margy that she's not available, Emma makes a counter proposal.

```
(16) NB IV:9
01 Emm:     Honey I'll come down after I had muh liddle bowl a'soup'n
02          salad'n I'll call'em ba:ck to yuh I'd love it.
03          (1025 ms)
04 Mar:     We:ll (0.7) Oka:y [I:-uh: (.) I wanteda (j's)
05 Emm:                       [D'you haftuh have it done no:w?
```

The first component of Margy's response is the prefatory particle "well", which our own analysis shows is strongly associated with dispreferred actions (see The Frequency of Turn-Initial Practices). After a long pause, within which Emma could potentially interject to revise her proposal, Margy responds with "Oka:y", produced with a relatively low pitch and compressed pitch range, and thereby accepts the proposal, albeit in a hearably begrudging manner. Emma is apparently not deaf to this. In a position in which she could bring the sequence to a close (e.g., with a sequence-closing third), she expands the sequence with a question that orients to the acceptance as less than adequate (line 5). Although Margy accepts

Emma's proposal, she does so with a turn format that qualifies her commitment to that very acceptance.

To investigate the possibility that the timing of a response may be a feature of a turn's format, not its action per se, a distinction was made between preferred actions that were qualified in some way and those that were not. For a response to be considered qualified, it had to satisfy one or more of the following criteria:

1. The acceptance was conditional on a date or time, a third party's acceptance, or an external contingency (e.g., whether the speaker will have access to a phone).
2. The phonetic realization of the acceptance displayed a negative affective stance (e.g., *sotto voce*, low pitch, compressed pitch range).
3. The recipient of the acceptance subsequently treated it as insufficient.
4. The acceptance included linguistic forms that explicitly qualify the speaker's commitment (e.g., "I think", "maybe").

The analysis was based exclusively on the criteria listed above and did not take into account the timing of the response or the presence of turn-initial particles. Just over one quarter of preferred actions met these criteria.[5] To assess the reliability of this analysis, a research assistant with no prior experience in CA coded 20 percent of the preferred responses using the criteria above. The vast majority of cases (92%) received the same value as in our own coding, yielding a very high level of agreement in a Cohen's kappa coefficient ($\kappa = 0.83$).

---

[5] Schegloff (2007) notes that one common format for a dispreferred action is a pro-forma agreement followed by a disagreement, the paradigm case of which is a "yes, but..." response (p. 69-70; cf. Sacks 1973/1987). We do not use this terminology because we find that the acceptance components that occur in qualified acceptances are rarely 'pro forma'. In the majority of cases, these components display some degree of commitment to the future course of action, even as a subsequent turn component works to modify its terms.

Flat Rejections. Many dispreferred actions in the data include turn components such as prefaces ("well", "uhm"), qualifications ("maybe", "I think"), and terms of endearment ("honey", "sweetie") that in various ways work to minimize the impact of a response that fails to align with the prior speaker's action. But the data also include many dispreferred actions that lack these components and thereby employ a turn format that formally resembles a typical preferred action. In this section, we exemplify a range of such cases in which dispreferred actions are delivered through preferred turn formats and outline the formal criteria that we use to identify these *flat rejections* within the data.

The formal similarity between some preferred and dispreferred actions is evident in the following cases, an acceptance and a rejection, respectively.

```
(17) Holt U88:2:1
01 Les:     So: uh we ↑wondered if perhaps we'd give that
02          a try: what d'you thin:k.
03          (098 ms)
04 Arn:     What a good idea:.
(18) 3b A&B3
01 Bel:     Well we'll go somewhere else:.
02          (269 ms)
03 Anne:    No no no no.=I don't want you to eh:
04          (0.5)
```

In the first case, Arnold accepts Leslie's proposal with a relatively simple turn format that lacks prefatory particles that would delay the beginning of the base turn-constructional unit and exhibits no mitigation or qualification. Similarly, in the second case, after Bel offers, for the first time, to change the terms of a previous invitation, Anne rejects this with a turn format that has a rejection component in turn-initial position and displays no mitigation or qualification. Thus a dispreferred action, rejection, can be delivered through a preferred turn format, that is, one that lacks components that qualify or mitigate the impact of the action.

A similar turn design can be seen in the response that Nan gives to Mark's invitation in the following example.

```
(19) Holt SO88:II:2:6
13 Nan:     I ↑can't get mu:ch what you said you said carnival,
14 Mar:     ehYeh the carnival toni:ght.=
15 Nan:     =Ye:s,
16 Mar:     .hh Do you want to go:
17          (106 ms)
18 Nan:     Oh no-:
19          (0.3)
20 Nan:     I'm too ti-:red Mark
```

After Nan first fails to hear Mark's announcement about a local carnival (lines 7-13), Mark reissues it and successfully secures a receipt (line 15) before he produces the invitation (line 16). The response that Nan provides has a simple design, lacking the turn components that often occur with dispreferred actions. Indeed the response has only two components: an "oh"-preface and "no" as an explicit rejection component. Note that the account Nan provides at line 20 comes after a gap at line 19 and occupies a separate turn-constructional unit from the rejection. The first component of the response, an "oh"-preface, is according to Heritage (1998) "a practice through which a speaker indicates a problem about a question's relevance, appropriateness, or presuppositions" (p. 295). In this case, the "oh"-prefaced rejection treats even the possibility that the speaker would want to go to the carnival as problematic and hints that the inviter should have known better than to ask.

The dispreferred action in the next case, in which Joy rejects an offer by Leslie, has a more complex design but also lack the specific turn components associated with dispreferreds.

```
(20) Holt 5:88:1:2
01 Les:     .hhh ↑P'APS you'd like- Would ↑you like eh:m:: some
```

02 frozen f::::↓ruit fr'm our k- ou:r ↓freezer as a small
03 recompens[e?
04 **(-023 ms)**
05 Joy: [Oh: Les for goodness sake n:no I don't
06 Joy: want anything

The offer that Leslie makes emerges out of a gratitude-acceptance sequence and is explicitly formulated as compensation ("a small recompense") for a favor done by Joy. In this environment, an acceptance of the offer could implicate a selfish motive for what would otherwise be seen as altruistic behavior. Perhaps to combat this implication, Joy employs a battery of practices in the design of her response that work to render her rejection absolute. The "oh"-preface indexes a stance that the offer had not been expected, and thereby treats its relevance, tied to Leslie's display of gratitude, as problematic (cf. Heritage, 1998). The address term and "for goodness sake" register the offer as a mild offense (cf. Clayman, 2013). And the account ("I don't want anything") openly contradicts the implication that she might want something in return for the favor and that her actions could therefore have a selfish motive.

To investigate the relationship between the timing of a response and its format, we examined all dispreferred actions to identify those that (a) include "no" or other explicit disconfirmations and (b) do not include components that qualify or mitigate the action. The timing of the response was not taken into consideration. To assess the reliability of this coding, a research assistant with no prior experience in CA coded 20 percent of the dispreferred responses using these criteria. All of the cases (100%) received the same coding as in our own analysis.

The Timing of Qualified Acceptances and Flat Rejections. The analysis resulted in four response types: qualified acceptances (*n*=36, 18.5%) and flat rejections (*n*=18, 9.2%), together with the remainder of preferred and dispreferred actions, which we refer to as normal acceptances (n=87, 44.6%) and normal rejections (n=54, 27.7%). The density plots in Figure 5 show the timing of all four response types for Offset 2.

A comparison of normal and qualified acceptances shows that a significant portion of late preferred actions in the previous section were qualified acceptances. The mode of the distribution of normal acceptances is approximately 275 ms, whereas the mode for qualified acceptances is approximately 500 ms. Before 100 ms, over one quarter of normal acceptances occur, but qualified acceptances are quite rare (28.7% of normal acceptances vs. 5.5% of qualified acceptances; $\chi^2(1) = 6.69$, $p < .01$). Conversely, after approximately 700 ms, qualified acceptances are much more frequent than normal acceptances (25% and 2.3% respectively; $\chi^2(1) = 13.45$, $p < .0005$). That the timing of normal and qualified acceptances differs systematically is further confirmed by a mixed-effects regression model with Offset 2 as the dependent variable, acceptance type as a fixed factor and speaker as a random factor ($\beta = 342.6$, $t = 3.49$, $p < .0005$).

For normal and flat rejections, a comparison yields similar results. Flat rejections occur significantly earlier than normal rejections, with modes of approximately -50 and 325 ms, respectively. While few normal rejections occur before 0 ms, a considerable number of flat rejections occur in overlap with the prior turn (1.8% vs. 33.3%, respectively; $\chi^2(1) = 11.86$, $p < .001$). After 800 ms, there are no cases of flat rejection, but over one quarter of normal rejections (27.8%) occur after this point. A mixed-effects regression model with Offset 2 as the dependent variable, rejection type as a fixed factor and speaker as a random factor further confirmed that normal rejections occur later than flat rejections ($\beta = 630.9$, $t = 5.05$, $p < .0001$). These observations suggest that, like qualified acceptances, the timing of flat rejections differs systematically from their normal counterparts.

Interestingly, even after qualified acceptances and flat rejections have been separated out, the modes of the distributions for normal acceptances and normal rejections, which constitute the majority of preferred and dispreferred actions, are still quite similar, at 275 and 325 ms, respectively (cf. Figure 3, in which the modes were both approximately 250 ms). Only at the extremes of the distributions, however, can one observe clear differences. The proportion of responses that occur in overlap is much higher for normal acceptances than for normal rejections (14.9% vs. 1.9%; $\chi^2(1) = 5.01$, $p < .05$), and while close to one third of responses that occur after 700 ms are normal rejections, very few normal acceptances occur after this point (31.5% vs. 2.3%; $\chi^2(1) = 21.89$, $p < .0001$). The bulk of normal acceptances and rejections, however, still occur within the same temporal window, between 0 and 700 ms.

These results show that the details of a turn's construction (e.g., whether or not the turn includes components that qualify the speaker's commitment) is one factor that influences the timing of a responding turn, and one that does so more decisively than whether the turn delivers a preferred or dispreferred action. This suggests that the timing of a turn is linked not to the action of the response per se (i.e., acceptance or rejection) but rather to the specific practices that speakers use to construct the action.

This can be seen clearly in the density plots in Figure 6, in which the normal acceptances and flat rejections are combined as preferred formats and normal rejections and qualified acceptances are combined as dispreferred formats (for Offset 2). In this analysis, 53.8 percent of cases (*n*=105) have preferred turn formats and 46.2 percent (*n*=90) have dispreferred turn formats. In comparison to the distributions of preferred and dispreferred actions for the same measure in Figure 3, one can observe a systematic difference in timing between preferred and dispreferred formats, not only at the extremes of the distributions but also in their cores. The mode of the distribution of preferred formats is between 175 and 275 ms, whereas the mode of dispreferred formats is approximately 375 ms. This difference of over 100 ms stands in stark contrast to the near identical modes of preferred and dispreferred actions (see Figure 3). Further inspection of the data revealed that preferreds were more common under 300 ms (71.6% of responses between 0 and 150 ms and 76.1% of those between 150 and 300 ms were preferred). However, after 300 ms the proportion of preferreds drops drastically to 41.6%. This drop continues as the duration of the gap increases (23.9% between 450 and 600 ms, and 15% after 600 ms). A mixed effects regression model with Offset 2 as dependent variable, format as fixed predictor, and speaker as random factor, confirms these observations that dispreferred formats tend to have later timings than preferred formats ($\beta = 410.44$, $t = 5.78$, $p < .0001$).

The differences in the modes of the two distributions, as well as the drastic change in the proportion of dispreferred formats after 300 ms, leads us to suggest that timing is more strongly associated with turn formats than with preferred and dispreferred actions per se, for which we observed similar modes and a higher degree of overlap in the distributions.

Are Objective Measurements of Timing a Problem?

The standard method for the timing of turn-taking in CA does not use a computer to measure the duration of gaps and pauses, as we have done, but rather employs a ‘counting phrase’ (e.g., ‘one Mississippi, two Mississippi,’ and so on), the pace of which the analyst adapts to the pace of the preceding talk. The more of the phrase the analysts can produce within the silence, the longer the duration of the silence is said to be (see Hepburn & Bolden, 2013). In this way, the standard method aims to produce relative measurements of the timing of gaps and pauses that take into account a variety of factors, such as speech rhythm and tempo, that may influence the perception of timing in conversation. To determine whether our results could be an artifact of our objective method of timing, we investigated (a) the relationship between gap duration and speech rate and (b) the relationship between our objective measurements and those of Gail Jefferson in her own transcriptions, which are available for the NB, SBL, and Holt corpora.

A first possibility that we consider is that the rate of speech of a first pair part may influence the amount of time a next speaker takes to produce a response. That is, a first pair part with a low speech rate might result in a longer gap, as the next speaker adjusts the timing of his or her response to that of the prior speaker. To investigate this possibility, we calculated the speech rate of first pair parts, measured as the number of syllables per second, and fitted a series of mixed-effects regression models with our different offsets as dependent variables, speech rate as a predictor, and speaker as a random factor. In none of these models did speech rate approach statistical significance (Offset 1: $\beta = 0.028$, $t = 0.15$; Offset 2: $\beta = 0.077$, $t = 0.44$; Offset 3: $\beta = 0.102$, $t = 0.85$). Therefore we conclude that the rate of speech of first pair parts and the timing of responses are not related and, more importantly, that the substantial variability and overlap in timing between preferred and dispreferred actions in our data cannot be attributed to variation in speech rate.

A second possible problem is that our objective measurements of timing may differ substantially from the relative measurements of timing typically used in CA. In order to check this, we compared our measurements to those of Gail Jefferson in her own transcriptions, which are available for the bulk of our data (76.9% of cases, n=150). Figure 7 illustrates the relationship between turn-initial silence (Offset 1) and Jefferson's timing. Jefferson's timing includes all cases which had a numeric value (which exhibited a range of [0.2, 1.2]), plus those annotated as 'no gap', 'latch', and 'micropause' (see Hepburn & Bolden, 2013 for discussion of these terms). In this figure, 'no gap' and 'latch' cases are displayed with a value of 0, whereas 'micropause' cases are shown with a value of 0.1. The solid line illustrates the fit of a linear model regressing Jefferson's timing on Offset 1. This model was statistically significant ($R^2 = .88$, $p < .0001$), and had an intercept of -121.1, and a slope of 0.839, indicating that Jefferson's timing undershoots objectively measured time by roughly 120 ms and that there is an additional undershoot of approximately 15 percent for each 1000 ms. Putting these details aside, the model shows that there is a strong correlation between Jefferson's timing and Offset 1. This correlation is further illustrated in Figure 8, which shows Jefferson's timing as a function of preference status. As in Figure 2, there is a substantial amount of overlap between preferreds and dispreferred actions in early responses and a bigger proportion of dispreferred actions in late responses. The considerable variability and overlap in the timing of preferred and dispreferred actions presented in the previous sections cannot be attributed to the use of an objective method of timing.

Discussion

Our investigation into the timing and construction of preference began with two goals. The first was to reproduce, with quantitative methods, the classic finding in CA that dispreferred responding actions tend to be delayed. The second and more central goal was to assess the claim that speakers can use the initial timing of response to anticipate its status as preferred or dispreferred and thereby anticipate its action (e.g., as acceptance or not). In this section, we review and discuss the results of our study in light of these goals.

Are Dispreferred Actions Delayed?

A common generalization in the CA literature is that dispreferred responding actions tend to be delayed relative to preferreds (see Background). So stated, this claim is too vague for straightforward verification. The most precise formulation in the literature comes from Schegloff (2007), who observes that "[t]he transition space between the first pair part turn and a dispreferred second pair part turn is commonly overlong" (p. 67). Although this formulation is technically precise in some respects, Schegloff does not specify what constitutes a "common" occurrence nor what counts as an "overlong" transition space. The results of our study indicate that one's assessment of the generalizations in the literature depends crucially on the specification of details such as these. In particular, our study shows that the answer depends on (i) the method that one uses to measure the gaps between turns,

(ii) whether one considers the most frequent or the most extreme cases, and (iii) whether one analyzes preference as an action-based or format-based phenomenon. We discuss each of these points in turn.

Research on the timing of turn-taking generally treats the gap between two turns at talk as a simple phenomenon with two discrete boundaries. The timing of a transition between speakers is the duration, in milliseconds, of the gap between the end of a first turn and the first component of a second turn, including in-breaths (see, e.g., Stivers & Enfield, 2010, p. 2625). This conservative measure of timing considers only the duration of silence between turns and equates to our Offset 1. Under this measure, we find no tendency for dispreferred actions to be delayed. Indeed, due to the greater frequency with which dispreferreds occur with turn-initial in-breaths and the freedom for in-breaths to occur in overlap with the prior turn, the most frequent dispreferred actions occur approximately 50 ms earlier than preferreds for Offset 1. A measure of timing that corrects for this, including preparatory actions such as in-breaths and clicks within the measure of the transition space, also fails to reveal a tendency for dispreferred actions to be delayed. Under this measure, Offset 2, preferred and dispreferred actions both occur after roughly 250 ms, a duration that constitutes a ‘normal’ transition in the organization of turn-taking (Jefferson, 1984, p. 18; Schegloff, 2000, p. 51-52; Stivers et al., 2009). In contrast, we observe a decisive difference of approximately 300 ms in the timing of preferred and dispreferred actions under a measure of timing, Offset 3, that includes not only preparatory actions but also prefatory particles such as “well” and “u(h)m” within the transition time. This corroborates observations made by Atkinson and Drew (1979), Heritage (1984), and Pomerantz (1984), among others, that turn-initial practices such as these work to delay dispreferred actions. But Offset 3 is not a standard measure of the transition space, insofar as it includes turn-initial components that constitute the beginning of the responding turn. Thus we find no evidence that the most frequent dispreferred responding actions tend to occur after longer transition spaces than the most frequent preferreds.

An analysis of the most extreme cases, however, points to a different conclusion. After approximately 700 ms, the proportion of dispreferred actions is significantly greater than that of preferreds (for Offsets 1 and 2). Therefore, although preferred and dispreferred actions most frequently occur with little or no delay, dispreferred actions are more common than preferreds after relatively long delays. This observation explains how an analysis of timing that reports mean gap durations (e.g., Stivers et al., 2009), but not the complete distributions as we have done, could conclude that the timing of two response alternatives differs even if the timing of the most frequent cases may in fact be the same. Insofar as long gaps are systematically more frequent than long overlaps in conversation, the distribution of gap durations in conversation will be non-normal. As a measure of central tendency, the mean is notoriously sensitive to extreme cases and thus may not be appropriate for the study of turn-taking (Heldner & Edlund, 2010, p. 557). Therefore the observation that the proportion of dispreferred actions increases after approximately 700 ms suggests that late dispreferred actions would disproportionately influence the mean gap duration for dispreferreds and could thereby obscure the fact that the most frequent dispreferreds occur with little or no delay.

Our results also cast doubt on the association between dispreferred actions and relatively short delays. Although relatively short gaps have been associated with dispreferred actions in the literature (e.g., Schegloff, 2007, p. 68), this has not always been the case. In a classic study, Pomerantz (1984) supports the claim that “a conversant, in the course of producing a disagreement, may initially respond with silence” with an array of four dispreferred responding actions, all of which occur after relatively long gaps of 600 ms or more (p. 70-71). Our results suggest that such cases in fact represent the true phenomenon, namely that *long* gaps are associated with disagreement and rejection. Relatively short gaps are no more associated with dispreferred actions than preferreds. In practical terms, the quantitative evidence thus indicates that the analysis of a response as a dispreferred action cannot rest on the observation of a short delay in production.

In a format-based analysis of preference, however, even relatively short delays are associated with dispreferred turn formats, be they forms of acceptance or rejection. Given that preference is both an action-based and format-based phenomenon (Schegloff, 1988), the generalization in the literature that dispreferred responses tend to be delayed can be understood in two ways, either as a claim about actions per se or as a claim about the formats that speakers use to deliver actions. Under a format-based analysis, our results demonstrate a systematic difference in the timing of preferred and dispreferred turn formats, one that affects not only the most extreme cases but also the most frequent ones. Whereas the most frequent preferred and dispreferred actions both occur within the same temporal window centered around 250 ms, the most frequent dispreferred formats occur approximately 100 ms later than the most frequent preferred formats. Furthermore, after only 300 ms, the proportion of dispreferred formats becomes greater than that of preferred formats, a difference that increases even more substantially after 600 ms.

The Reliability of Timing as an Early Signal of Preference

A generalization in the literature is that the timing of a responding turn can enable a recipient to project whether the response will be a preferred or dispreferred action, such that turn-initial silence “may be interpreted as the first move toward some form of disagreement/rejection” (Clayman, 2002, p. 235). While our results support the general conclusion that the timing of a responding turn can facilitate early projection, they also suggest that the nature and reliability of this projection differs as the duration of the transition space increases.

The statistical trends we observe suggest that a recipient who hears a response begin in overlap or after a normal transition space of 0 to 300 ms could use information about timing to inform a prediction that the response should be a preferred action in a preferred format, which occurs in 62.9 percent of cases in this window, as opposed to the next most common response type, dispreferred actions in dispreferred formats, which occur in 19.2 percent of cases. After approximately 300 ms, however, the prediction should change. A recipient who hears more than 300 ms of silence could use this to predict that the action should be preferred, but at this point the recipient would also predict that it should have a dispreferred turn shape, one which qualifies the speaker's commitment to the course of action. That is, a gap of more than 300 ms would be sufficient to project that a straightforward acceptance (i.e., a preferred action in a preferred format) is less probable, though still possible. After a relatively long gap of 700 to 800 ms or more, the prediction should change yet again. At this point, the proportion of dispreferred actions becomes greater than that of preferreds and virtually all responses have a dispreferred turn format. A reasonable prediction would therefore be that the response should be a dispreferred action in a dispreferred format, though a qualified acceptance is still possible.

The observation that predictions of an incipient action should change as the duration of the gap increases is supported by results from experimental research (Roberts et al., 2006, 2011; Roberts and Francis, 2013). This research has established a relationship between the duration of gaps before acceptances and attributions that participants make about the 'willingness' of speakers to assent to the requests. Roberts and Francis investigated the possible existence of a temporal threshold at which participants' attributions of willingness change from neutral to negative, using scripted request-acceptance sequences, with gap durations manipulated to be from 200 to 1200 ms at 100 ms intervals. The authors observed no statistically significant differences in ratings between intervals from 200 to 700 ms, a temporal window in which the bulk of normal acceptances (the response type that most closely resembles the experimental stimuli) occur in our data (see Figure 5). Between 700 and 800 ms, however, the authors observed a statistically significant decrease in the rating of speakers' willingness to comply with requests. Our results indicate that only normal rejections and qualified acceptances frequently occur after 700 ms (see Figure 5). This could be viewed as converging evidence for a temporal threshold. Indeed, Roberts and Francis themselves note the connection between their results and early results from the current study (Kendrick and Torreira, 2012), pointing out that the significant decrease in ratings between 700 and 800 ms coincides with our observation that dispreferred responding actions only become more frequent than preferreds after approximately 700 ms.

Our results warrant two general conclusions. First, the timing of a response appears to be a variable signal. A small departure from a normal turn transition (i.e., a gap of between 300 and 700 ms) may alert the recipient that the most frequent response type, a normal acceptance, is less likely, whereas a relatively large departure (i.e., a gap of 700 to 800 ms or more) may be grounds for the recipient to infer that rejection is imminent. Second, insofar as we observe the strongest relationship between timing and turn formats, the timing of response is best understood as a turn-constructional feature, the first virtual component of a preferred or dispreferred turn format, one without a one-to-one relationship to the actions speakers use it to perform (Schegloff, 1984).

Acknowledgements

We would like to thank Elizabeth Holt and Jack Sidnell for access to data, and Steve Levinson for the initiation of this project and much subsequent input. This research was made possible by the financial support of the Language and Cognition Department at the Max Planck Institute for Psycholinguistics, Max-Planck Gesellschaft, and European Research Counsel's Advanced Grant 269484 "INTERACT" to Steve Levinson.

References

Atkinson, J. M., & Drew, P. (1979). *Order in court: The organisation of verbal interaction in judicial settings*. Atlantic Highlands, NJ: Oxford Socio-Legal Studies.

Clayman, S. E. (2002). Sequence and solidarity. In S. R. Thye & E. J. Lawler (Eds.), *Group Cohesion, Trust and Solidarity* (pp. 229–253). Oxford: Elsevier Science Ltd.

Clayman, S. E. (2013). Agency in response: The role of prefatory address terms. *Journal of Pragmatics, 57,* 290-302.

Couper-Kuhlen, E. (in press). What does grammar tell us about action? *Pragmatics*.

Curl, T. S. (2006). Offers of assistance: Constraints on syntactic design. *Journal of Pragmatics*, *38*(8), 1257–1280.

Curl, T. S., & Drew, P. (2008). Contingency and action: A comparison of two forms of requesting. *Research on Language and Social Interaction*, *41*(2), 129–153.

Davidson, J. (1984). Subsequent Versions of Invitations, Offers, Requests and Proposals, Dealing with Actual or Potential Rejection. In J. M. Atkinson & J. Heritage (Eds.), *Structures of Social Action: Studies in Conversation Analysis* (pp. 102–128). Cambridge: Cambridge University Press.

Dersley, I., & Wootton, A. (2000). Complaint Sequences Within Antagonistic Argument. *Research on Language and Social Interaction*, *33*(4), 375–406.

Drew, P. (1984). Speakers' reportings in invitation sequences. In J. M. Atkinson (Ed.), *Structures of social action: Studies in conversation analysis* (pp. 129–151). New York: Cambridge University Press.

Ford, C. E., & Thompson, S. A. (1996). Interactional units in conversation: Syntactic, intonational, and pragmatic resources for the management of turns. In E. Ochs, E. A. Schegloff, & S. A. Thompson (Eds.), *Interaction and Grammar* (pp. 134–184). Cambridge: Cambridge University Press.

Hayano, K. (2013). Question design in conversation. In J. Sidnell & T. Stivers (Eds.), *The Handbook of Conversation Analysis* (pp. 395–414). Malden: Blackwell Publishing Ltd.

Heldner, M., & Edlund, J. (2010). Pauses, gaps and overlaps in conversations. *Journal of Phonetics*, *38*(4), 555–568.

Hepburn, A., & Bolden, G. B. (2013). The conversation analytic approach to transcription. In J. Sidnell & T. Stivers (Eds.), *The Handbook of Conversation Analysis*. Blackwell Publishing Ltd.

Heritage, J. (1984). *Garfinkel and Ethnomethodology*. Cambridge: Polity Press.

Heritage, J. (1998). Oh-prefaced Responses to Inquiry. *Language in Society*, *27*(03), 291–334.

Houtkoop-Steenstra, H. (1987). *Establishing agreement: An analysis of proposal-acceptance sequences*. Dordrecht: De Gruyter.

Houtkoop-Steenstra, H. (1990). Accounting for proposals. *Journal of Pragmatics*, *14*(1), 111–124.

Indefrey, P., & Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition*, *92*(1–2), 101–144.

Jefferson, G. (1984). Notes on some orderlinesses of overlap onset. In V. D'Urso & P. Leonardi (Eds.), *Discourse analysis and natural rhetoric* (pp. 11–38). Padua, Italy: Cleup Editore.

Kendrick, K. H., & Torreira, F. (2012). The timing and construction of preference: A quantitative study. Presented at the Discourse, Communication, and Conversation Conference, Loughborough University, Loughborough, England, UK.

Kitzinger, C., & Frith, H. (1999). Just Say No? The Use of Conversation Analysis in Developing a Feminist Perspective on Sexual Refusal. *Discourse & Society*, *10*(3), 293–316.

Lee, S. (2013). Response Design in Conversation. In J. Sidnell & T. Stivers (Eds.), *The Handbook of Conversation Analysis* (pp. 415–432). Malden: Blackwell Publishing Ltd.

Lerner, G. H. (1996). Finding “face” in the preference structures of talk-in-interaction. *Social Psychology Quarterly*, *59*(4), 303–321.

Levinson, S. C. (1983). *Pragmatics*. Cambridge: Cambridge University Press.

Levinson, S. C. (2013). Action formation and ascription. In J. Sidnell & T. Stivers (Eds.), *The Handbook of Conversation Analysis* (pp. 101–130). Malden: Blackwell Publishing Ltd.

Lindström, A., & Sorjonen, M.-L. (2013). Affiliation in Conversation. In J. Sidnell & T. Stivers (Eds.), *The Handbook of Conversation Analysis* (pp. 350–369). Malden: Blackwell Publishing Ltd.

Pomerantz, A. (1978). Compliment Responses: Notes on the Co-Operation of Multiple Constraints. In J. Schenkein (Ed.), *Studies in the Organization of Conversational Interaction* (pp. 79–112).

Pomerantz, A. (1984). Agreeing and disagreeing with assessments: Some features of preferred/dispreferred turn shapes. In J. M. Atkinson & J. Heritage (Eds.), *Structures of Social Action: Studies in Conversation Analysis* (pp. 57–101). Cambridge: Cambridge University Press.

Pomerantz, A., & Heritage, J. (2013). Preference. In J. Sidnell & T. Stivers (Eds.), *Handbook of Conversation Analysis* (pp. 210–228). Cambridge: Cambridge University Press.

Raymond, G. (2000). *The structure of responding: Type-conforming and nonconforming responses to yes/no type interrogatives* (Unpublished Ph.D. dissertation). University of California, Los Angeles.

Raymond, G. (2003). Grammar and Social Organization: Yes/no Interrogatives and the Structure of Responding. *American Sociological Review*, *68*(6), 939–967.

Roberts, F., & Francis, A. L. (2013). Identifying a temporal threshold of tolerance for silent gaps after requests. *The Journal of the Acoustical Society of America*, *133*(6).

Roberts, F., Francis, A. L., & Morgan, M. (2006). The interaction of inter-turn silence with prosodic cues in listener perceptions of “trouble” in conversation. *Speech Communication*, *48*(9), 1079–1093.

Roberts, F., Margutti, P., & Takano, S. (2011). Judgments Concerning the Valence of Inter-Turn Silence Across Speakers of American English, Italian, and Japanese. *Discourse Processes*, *48*(5), 331–354.

Robinson, J. D., & Bolden, G. B. (2010). Preference organization of sequence-initiating actions: The case of explicit account solicitations. *Discourse Studies*, *12*(4), 501 –533.

Sacks, H. (1987). On the preferences for agreement and contiguity in sequences in conversation. In G. Button, & J. R. E. Lee (Eds.), *Talk and Social Organization* (pp. 54–69). Clevedon, UK: Multilingual Matters.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language*, *50*(4), 696–735.

Schegloff, E. A. (1984). On Some Questions and Ambiguities in Conversation. In *Structures of Social Action* (pp. 266–298). Cambridge: Cambridge University Press.

Schegloff, E. A. (1988). On an Actual Virtual Servo-Mechanism for Guessing Bad News: A Single Case Conjecture. *Social Problems*, *35*(4), 442–457.

Schegloff, E. A. (1996). Turn organization: One direction for inquiry into grammar and interaction. In E. Ochs, E. A. Schegloff, & S. A. Thompson (Eds.), *Interaction and Grammar* (pp. 52–133). Cambridge: Cambridge University Press.

Schegloff, E. A. (2000). Overlapping Talk and the Organization of Turn-Taking for Conversation. *Language in Society*, *29*(01), 1–63.

Schegloff, E. A. (2007). *Sequence Organization in Interaction: A Primer in Conversation Analysis*. Cambridge: Cambridge University Press.

Schegloff, E. A., & Sacks, H. (1973). Opening up Closings. *Semiotica*, *8*(4).

Speer, S. A. (2012). The Interactional Organization of Self-Praise Epistemics, Preference Organization, and Implications for Identity Research. *Social Psychology Quarterly*, *75*(1), 52–79.

Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., … others. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences*, *106*(26), 10587.

Stivers, T., & Robinson, J. D. (2006). A preference for progressivity in interaction. *Language in society*, *35*(03), 367–392.

Stivers, Tanya. (2010). An overview of the question-response system in American English conversation. *Journal of Pragmatics*, *42*(10), 2772–2781.

Stivers, Tanya, & Enfield, N. J. (2010). A coding scheme for question–response sequences in conversation. *Journal of Pragmatics*, *42*(10), 2620–2626.

Tables

Table 1

The four response types investigated in this section

	Preferred format	Dispreferred format
Preferred action	normal acceptance	qualified acceptance
Dispreferred action	flat rejection	normal rejection

Figures

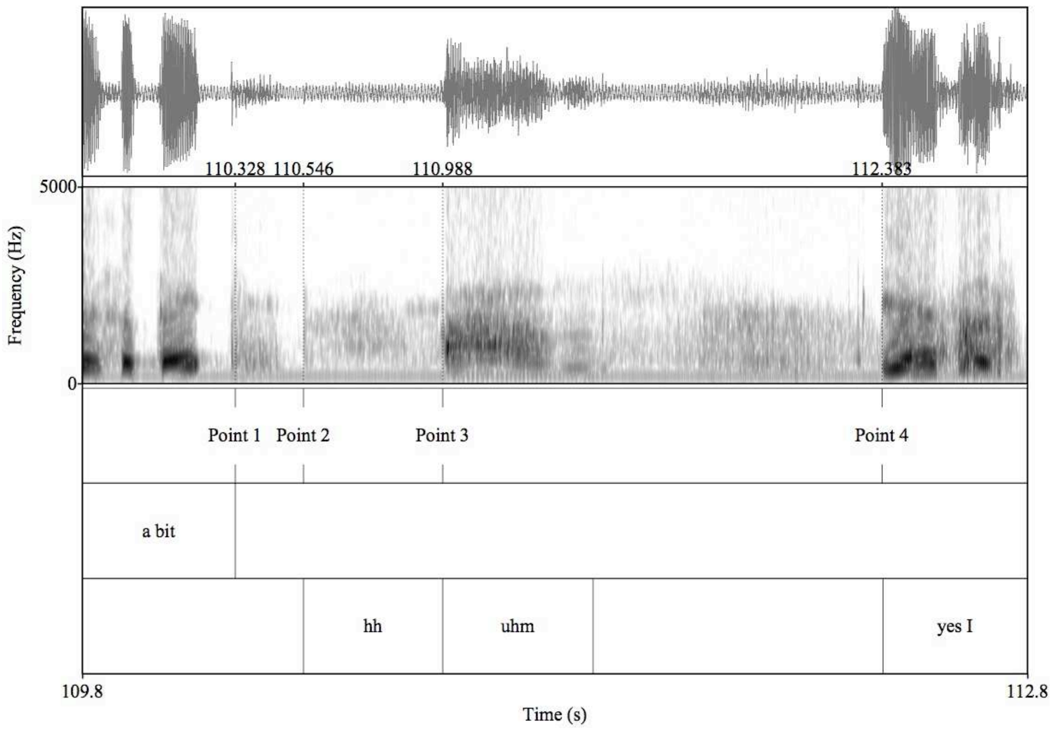

Figure 1. Location of the four temporal points in our measurement scheme as applied to the sequence Holt O88:1:9 105.778.

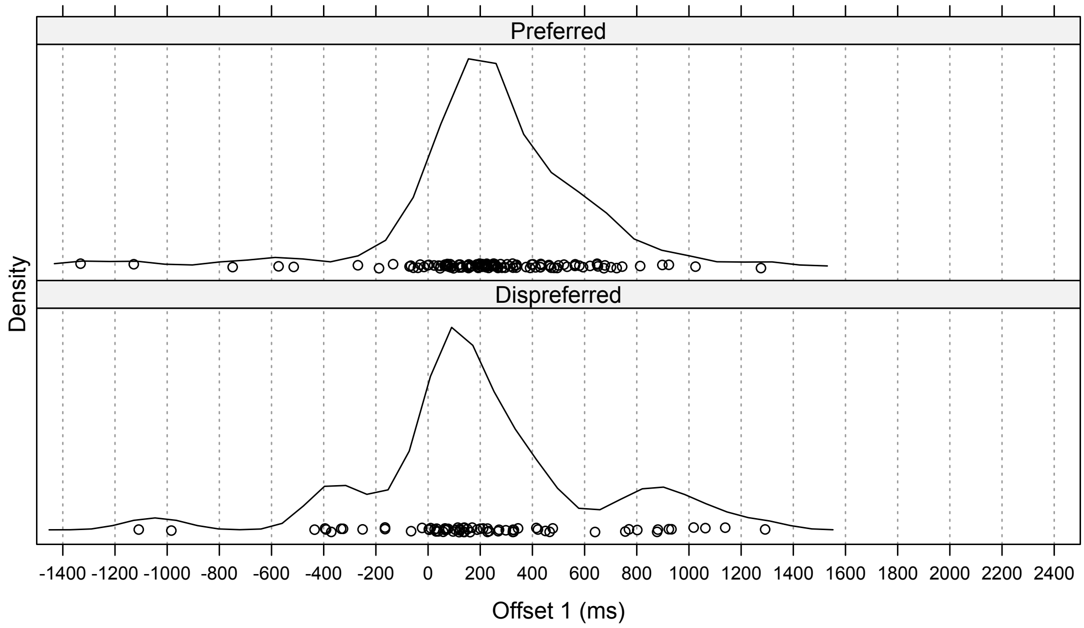

Figure 2. The timing of the first audible component of the response (Offset 1) in preferred and dispreferred actions.

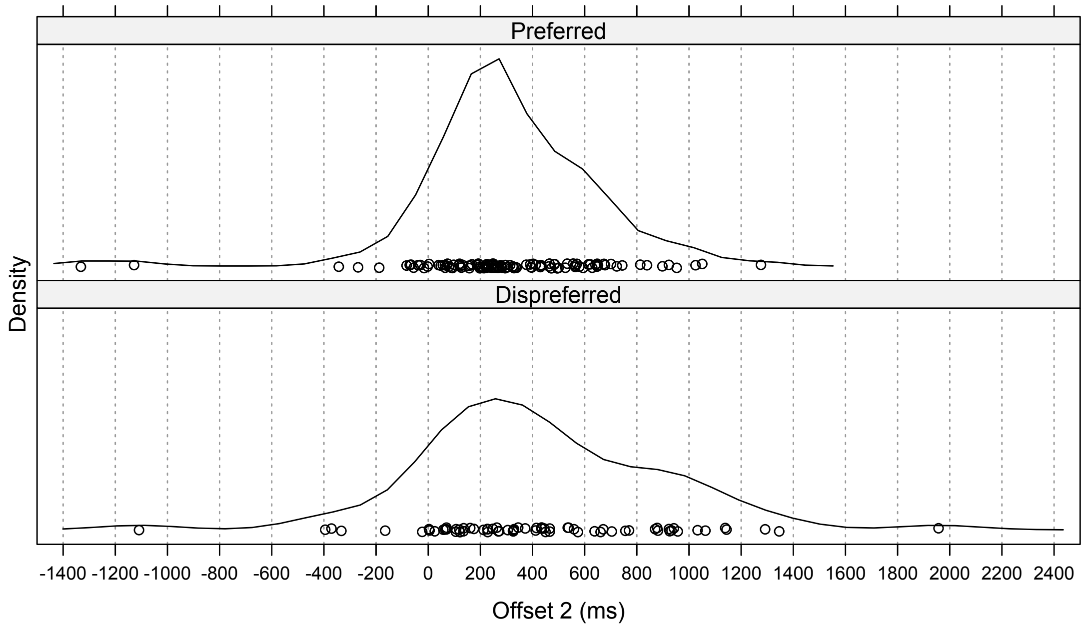

Figure 3. The timing of the first word or particle of the response (Offset 2) in preferred and dispreferred actions.

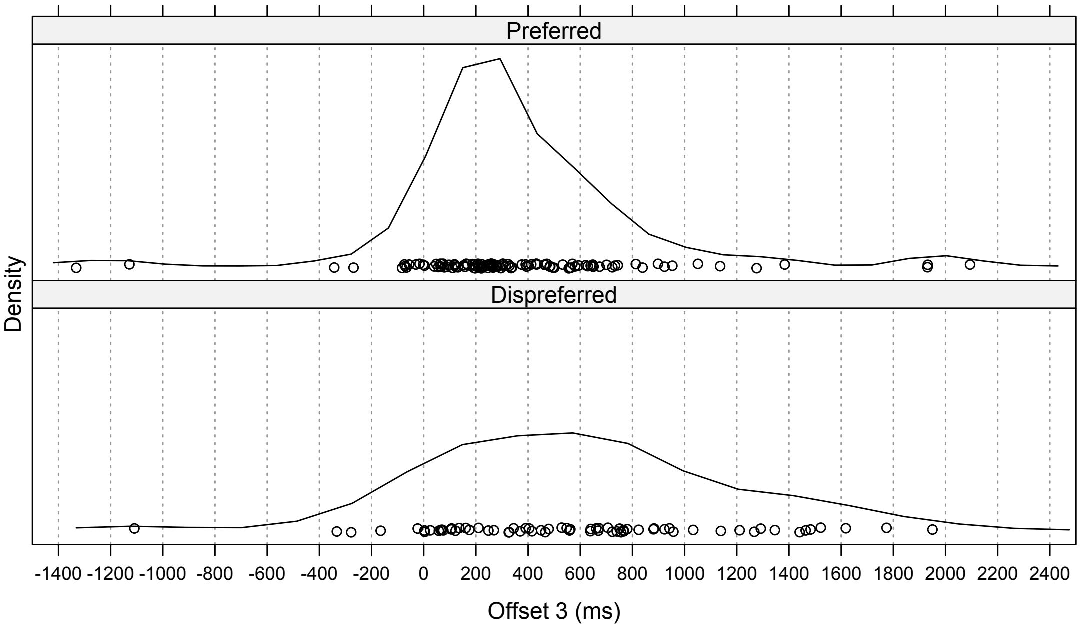

Figure 4. The timing of the turn proper (Offset 3) in preferred and dispreferred actions.

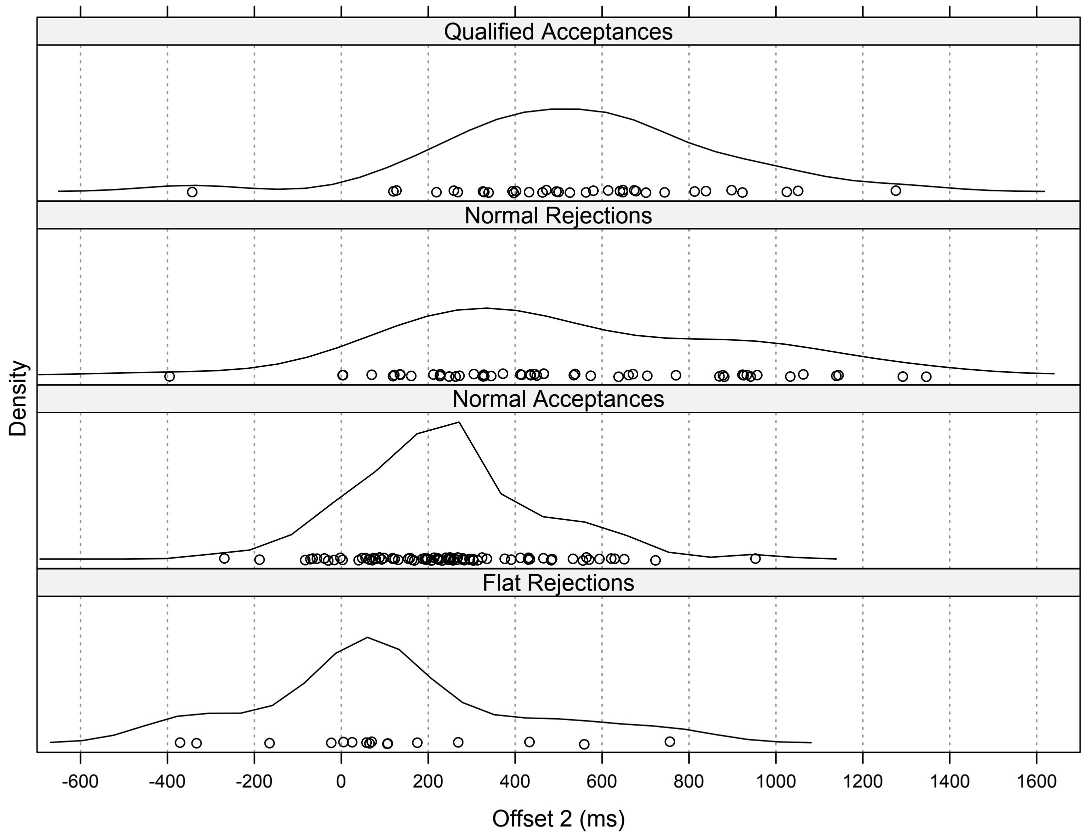

Figure 5. The timing of the first word or particle of the response (Offset 2) in normal acceptances, qualified acceptances, flat rejections, and normal rejections.

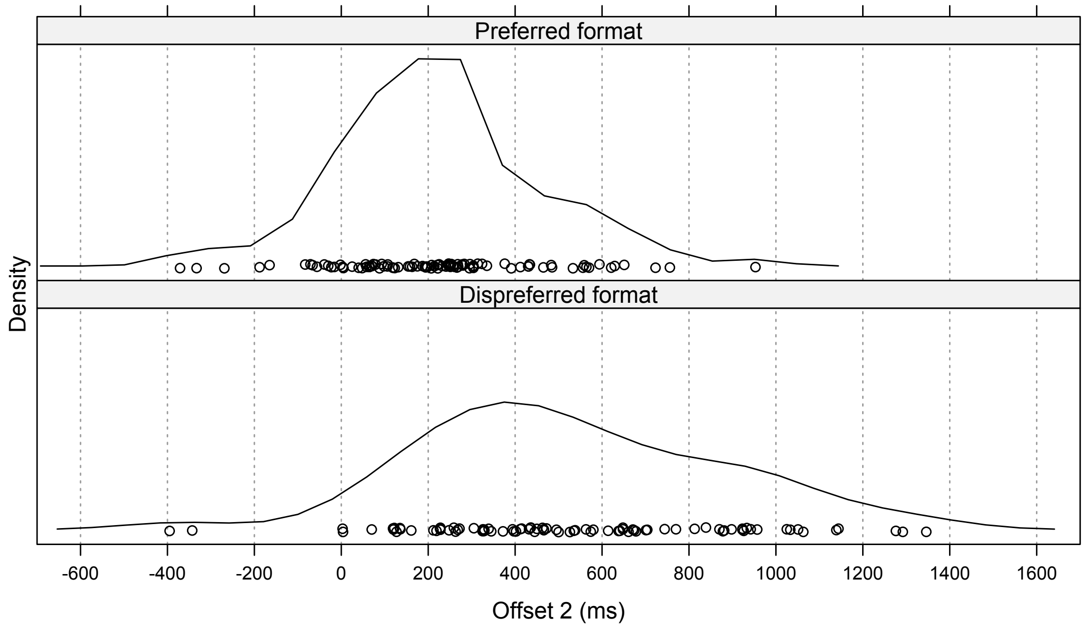

Figure 6. The timing of the first word or particle of the response (Offset 2) for preferred and dispreferred turn formats.

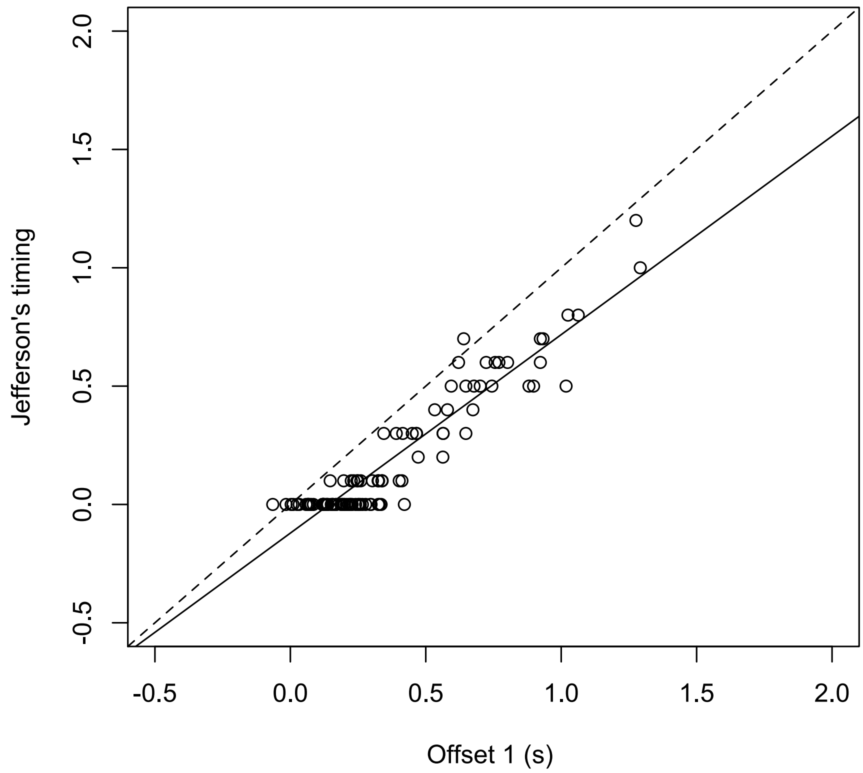

Figure 7. Jefferson's timing annotations (see text for details) as a function of the timing of the first audible component of the response (Offset 1). Non-numerical values in Jefferson's annotations (e.g., 'no gap') were adapted to values of 0 and 0.1 (see text for details). A least-squares regression line is shown in black, while a one-to-one relationship is shown with a dashed line.

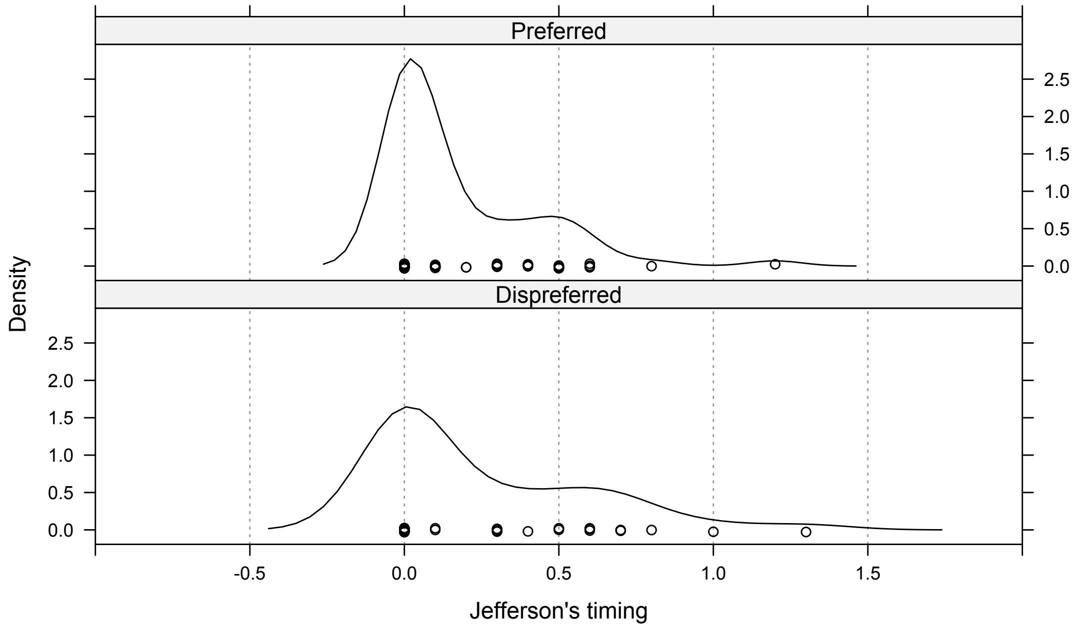

Figure 8. Jefferson’s timings for preferred and dispreferred actions.